Int.J.Curr.Microbiol.App.Sci (2017) *6*(12): 1327-1333

EXCELLENT PUBLISHERS

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 Volume 6 Number 12 (2017) pp. 1327-1333
Journal homepage: http://www.ijcmas.com

IJCMAS INTERNATIONAL JOURNAL OF CURRENT MICROBIOLOGY AND APPLIED SCIENCES 2017 DECEMBER www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2017.612.150

Identification of Potential Reference miRNA for qRT-PCR Studies in Cancers Using miRNA-seq Data

Rajesh Kumar, Avinash Marwal and R.K. Gaur*

Department of Biosciences, College of Arts, Science and Humanities, Mody University, Lakshmangarh, Sikar – 332311, Rajasthan, India
**Corresponding author*

Keywords

TCGA, qRT-PCR, Reference miRNA cancer.

Article Info

Accepted:
12 October 2017
Available Online:
10 December 2017

A B S T R A C T

Quantitative real-time polymerase chain reaction (qRT-PCR) is a low cost, rapid, accurate method for quantification of gene expression. Accuracy of qRT-PCR expression quantification is highly dependable on selected reference genes or miRNAs as these genes or miRNAs work as internal controls for normalization to quantify expression among different samples. Problems have been reported with steady expression or stability of reference genes in various conditions. Thus suitable reference genes or miRNAs need to be identified in various experimental conditions. The Cancer Genome Atlas (TCGA) Level 3 miRNA expression data for 11 cancers were collected, after pre-processing data from 4 cancers having 102 normal samples and 1428 tumor samples was used to find potential reference miRNA. Mean, coefficient of variation among normal and tumor samples were checked to find highly, stably expressed miRNAs between normal and tumor conditions. We found 32 potential qRT-PCR reference miRNAs in 3 out of 4 cancer types. 15 miRNA were found in cholangiocarcinoma (CHOL), 13 miRNA in kidney renal papillary cell carcinoma (KIRP) and 4 miRNA in stomach and esophageal carcinoma (STES). We did not find any miRNA passing all selection criteria in glioma (GBMLGG). In this study we used TCGA data to find potential qRT-PCR reference miRNAs in 4 cancers. We performed computational analysis of 1530 expression profiles of both normal and tumor conditions from 4 human cancer types and found total 32 stably and highly expressed potential reference miRNAs in 3 out of 4 cancer types. Data generated using next-generation sequencing (NGS) technologies can be very helpful in finding reference qRT-PCR miRNAs.

Introduction

qRT-PCR is a very popular variant of PCR which is used for low cost, rapid and accurate expression quantification of transcripts. Gene expression quantification using qRT-PCR is done with the help of reference genes. These genes are used as internal controls to normalize and remove biases introduced by differences in RNA content extraction, transfer, storage, quantity and quality of RNA material, reverse transcriptional activity across different samples (Bustin *et al.,* 2009). Reference genes are chosen for normalization based on their high and stable expression. Expression levels of reference genes don't vary much under different conditions, organisms, tissues and samples under study consideration. Most common reference genes are housekeeping genes such as GAPDH, 18s

rRNA, ACTB, GUSB, HMBS, TBP etc. (Kozera and Rapacz, 2013).

However ideally the expression of these references should be constitutively constantly high, and should not vary in different samples, nor be affected by experimental treatments, but in reality expression of the housekeeping genes or other most commonly used reference genes vary in different type of tissues, developmental stages, related species, abiotic stress, disease and infection, tumors and alternative splicing can also affect their expression (Glare *et al.,* 2002; Kozera and Rapacz, 2013). Thus identification of reference genes or miRNAs in various conditions is an important task in designing accurate and reliable qRT-PCR studies.

Apart from selecting housekeeping gene or other commonly used genes as reference there has been lot of attempts to identify reference gene in various cell lines and experimental conditions (Frericks and Esser, 2008; Kwon *et al.,* 2009; Lee *et al.,* 2007; Popovici *et al.,* 2009; Saviozzi *et al.,* 2006). Due to arrival of latest reference gene free expression quantification technologies such as microarray and NGS the data generated by these machines can be used for reference gene selection in desired tissue/condition.

Next generation sequencing technology variant such as RNA-seq has significantly reduced the cost and time of sequencing transcripts and is being used to generate gene expression data from various tissues and conditions.

This data is being made publicly available by submitting it to online databases such as Gene Expression Omnibus (GEO) (Barrett *et al.,* 2013), Sequence Read Archive (SRA) (Leinonen *et al.,* 2011), TCGA (http://cancergenome.nih.gov/) etc. TCGA is one specialized database in storing NGS data generated from various cancer tissues. We performed computational analysis of 1530 samples from 4 human cancer conditions and found total 32 stably and highly expressed potential reference miRNAs in 3 human cancer conditions. Further laboratory based validation will be required to find validity of these potential reference miRNAs.

Materials and Methods

miRNA expression data collection and preprocessing

TCGA Level 3 miRNA expression data generated using Illumina HiSeq platform for 11 cancers was collected from GDAC firehose website (https://gdac.broadinstitute.org/) in RPM format. GDAC firehose is a BROAD Institute's initiative to stores TCGA data in processed format. Cancers having 5 or more normal samples were included in the study.

From all the collected samples 7 cancer types did not have expression data for normal conditions. Finally we chose 4 cancer samples having 102 normal samples and 1428 tumor samples for analysis. Expression data of 5p, 3p arm of mature miRNA was taken in reads per million (RPM) format.

Selection of potential reference miRNAs

Total 4 criteria were set to select potential reference miRNAs. In this study, the definition of a reference miRNA is a miRNA that is highly and stably expressed across normal and tumor conditions. Potential reference miRNA for each cancer type were filtered using the following criteria, based on previous studies.

miRNA should express in all samples. Thus it should have RPM expression value > 0 in all samples.

Mean (N) > 100 and Mean (T) > 100.

CV (N) < 0.5 and CV (T) < 0.5.

Mean (N)/Mean (T) < 1.3 and Mean (T)/Mean (N) < 1.3.

Results and Discussion

To find potential qRT-PCR reference miRNA, we used miRNA-seq data from 4 cancer types more detail about number of total, tumor, normal samples are given in Table 1. Table 2 contains details about number of miRNAs passing various criteria for selection of reference miRNA as described in material and methods section. We found a total 32 reference miRNA satisfying the criteria we set to identify potential reference miRNAs in all 4 cancer types.

Cholangiocarcinoma (CHOL) had 15 miRNA, no miRNA passed selection criteria in glioma (GBMLGG), kidney renal papillary cell carcinoma (KIRP) 13 miRNA, stomach and esophageal carcinoma (STES) had 4 potential reference miRNA. Names of potential reference miRNA are given in Table 3.

Mean RPM expression values of various potential reference miRNA in different cancer conditions are plotted in Figure 1, 2 and 3. Many let-7 and hsa-mir-30 family miRNAs showed stable expression in various conditions. hsa-miR-361-5p was also stably expressed in CHOL, KIRP which was also discovered by Zhan and colleagues (Zhan *et al.,* 2014) in many other cancer types.

Most commonly used reference genes for qRT-PCR studies such as several housekeeping genes and non-coding RNA such as 18s rRNA are not suitable for expression quantification in all conditions due to inconsistency in their expression levels across tissue space and biological conditions (de Kok *et al.,* 2005). Gene sequences are much lengthier as compared to miRNA, making them unsuitable for miRNA qRT-PCR expression quantification studies. Thus Careful selection of an appropriate reference is extremely important as no single reference gene can serve as a universal control for all experimental conditions.

MicroRNAs are a class of small RNA molecules that regulate gene expression in many biological processes and disease conditions by mRNA degradation, translational repression, RNA-directed DNA methylation (Bartel, 2004; Floyd and Bowman, 2005; Millar and Waterhouse, 2005).

Reference miRNA for normalization and the miRNA being quantified should be from same type of sample for greater accuracy of expression quantification. Zhan *et al.,* (2014) used TCGA miRNA-seq data from 14 cancers having total 589 normal and 5727 tumor samples and identified total 126 potential reference miRNAs.

To extend their work and find suitable potential reference miRNA, we investigated TCGA high throughput miRNA-seq expression data from 4 cancer conditions. We identified 32 potential reference miRNA in 3 cancer conditions. miRNA such as let-7 family, hsa-miR-22 which showed highly stable expression is involved in various cancer suppression.

These miRNA are good for patient survival (Stahlhut Espinosa and Slack, 2006). Let-7 is known to be targeting tumorigenesis and angiogenesis and has also been reported as reference miRNA in other studies (Chang *et al.,* 2010).

Apart from these hsa-miR-361-5p was also reported by Zhan (Zhan *et al.,* 2014).

Table.1 Number of samples in each cancer tumor

Cancer type	Number of normal samples	Number of tumor samples	Total
Cholangiocarcinoma (CHOL)	9	36	45
Glioma (GBMLGG)	5	526	531
Kidney renal papillary cell carcinoma (KIRP)	34	292	326
Stomach and Esophageal carcinoma (STES)	54	574	628

Table.2 Number of miRNA in 4 cancer samples that passed our selection criteria as described in material and methods section

Cancer Type	miRNA having > 0 RPM in all samples	Normal samples		Tumor samples		Stable in both normal and tumor samples	Potential reference miRNA
		Highly Expressed	Stable	Highly Expressed	Stable		
Cholangiocarcinoma (CHOL)	2195	123	99	125	40	39	15
Glioma (GBMLGG)	1786	157	88	114	9	7	0
Kidney renal papillary cell carcinoma (KIRP)	2004	115	64	111	25	22	13
Stomach and Esophageal carcinoma (STES)	1725	116	29	127	14	9	4

Table.3 Potential reference miRNA identified in each cancer type

Cancer type	Potential reference miRNAs
Cholangiocarcinoma (CHOL)	hsa-let-7a-5p, hsa-let-7d-5p, hsa-let-7d-3p, hsa-let-7f-5p, hsa-miR-15a-5p, hsa-miR-16-5p, hsa-miR-26a-5p, hsa-miR-29a-3p, hsa-miR-103a-3p, hsa-miR-197-3p, hsa-miR-30c-5p, hsa-miR-145-5p, hsa-miR-186-5p, hsa-miR-361-5p, hsa-miR-151a-5p
Glioma (GBMLGG)	-
Kidney renal papillary cell carcinoma (KIRP)	hsa-let-7a-5p, hsa-let-7d-3p, hsa-miR-24-3p, hsa-miR-30a-5p, hsa-miR-92a-3p, hsa-miR-101-3p, hsa-miR-30d-5p, hsa-let-7g-5p, hsa-miR-23b-3p, hsa-miR-140-3p, hsa-miR-30e-5p, hsa-miR-361-5p, hsa-miR-151a-3p
Stomach and Esophageal carcinoma (STES)	hsa-miR-22-3p, hsa-miR-28-5p, hsa-miR-140-3p, hsa-miR-374a-3p

Fig.1 Mean RPM value plot of various potential reference miRNAs in cholangiocarcinoma

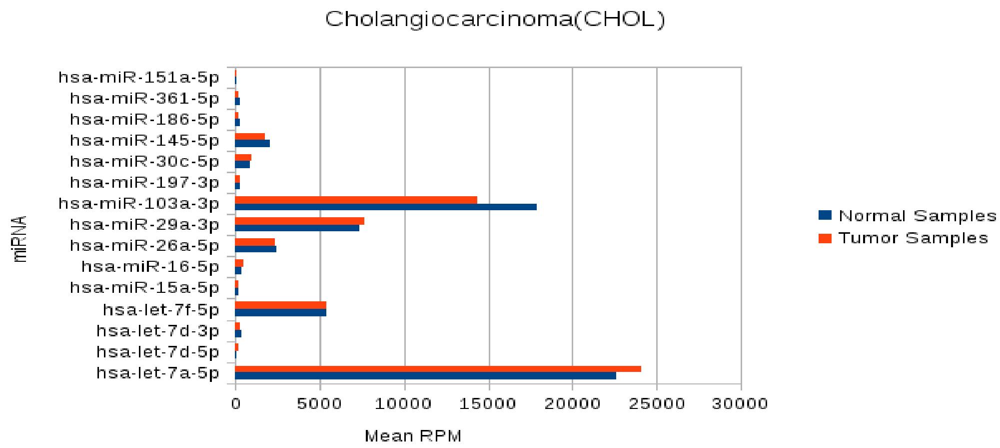


Fig.2 Mean RPM value plot of various potential reference miRNAs in kidney renal papillary cell carcinoma

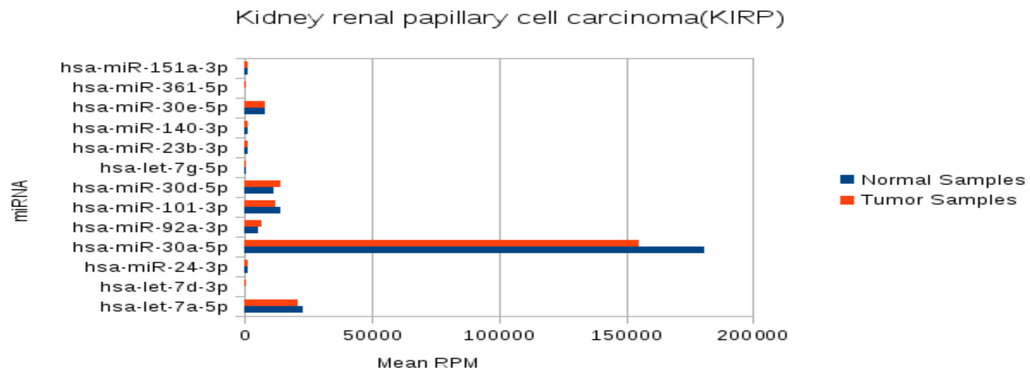
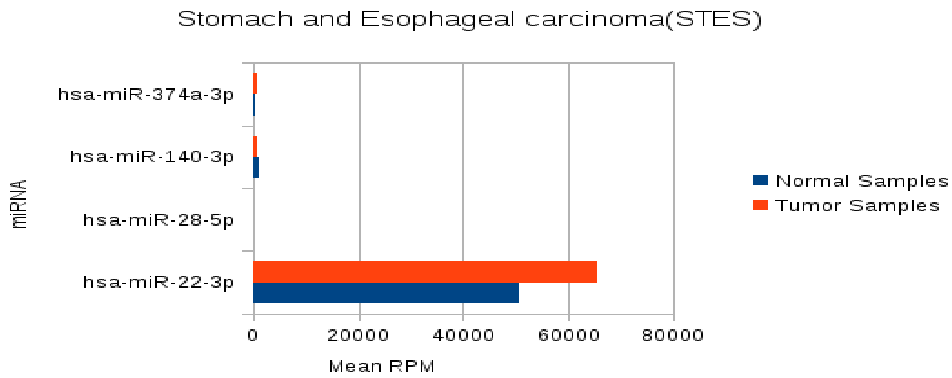


Fig.3 Mean RPM value plot of various potential reference miRNAs in stomach and esophageal carcinoma



Highly expressed (as compared to other potential qRT-PCR reference miRNA) both in normal and CHOL tumor cells, potential qRT-PCR reference miRNA 5p arm of hsa-let-7a which is known to modulate interleukin-6-dependent STAT-3 survival signaling in malignant human cholangiocytes (Meng *et al.,* 2007). Others were hsa-miR-103a-3p, hsa-let-7f-5p, hsa-miR-29a-3p. hsa-miR-30a-5p and has-miR-22-3p were highly expressed miRNA in KIRP and STES among other potential qRT-PCR reference miRNA. Further experimental validation is necessary to use these miRNA as control for expression quantification.

Most commonly used reference genes and miRNAs are not suitable as qRT-PCR reference miRNA because their expression can vary among various conditions. Thus identification of qRT-PCR reference miRNA becomes must to calculate expression using qRT-PCR. NGS technologies are reference free expression methods, thus can be utilized to find stably, highly expressed genes/miRNAs across samples. In this study we analyzed TCGA miRNA-seq data to find potential qRT-PCR reference miRNAs for 4 cancer conditions. Total 32 potential qRT-PCR reference miRNAs for 3 cancerous conditions were found. These miRNAs must be validated in wet lab before their use.

Acknowledgement

The authors are thankful to Department of Biotechnology, New Delhi, India for providing financial assistance through Bioinformatics National Certification (BINC) (File No. PU/BINC/2016/E-04), Science and Engineering Research Board – Department of Science and Technology, New Delhi, India for the financial assistance (File No. YSS/2015/000265). We would like to thank Mr. Purshotam Das for providing technical support while carrying out this research work.

Competing interests

Authors do not have any competing interests.

Author's contributions

Rajesh Kumar - Carried out the research work and wrote the manuscript.

Avinash Marwal - Drafting the article and revising it critically

R.K. Gaur - Corresponding author and final approval of the version to be published

References

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 41, D991–D995. doi:10.1093/nar/gks1193

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin. Chem. 55, 611–622. doi:10.1373/clinchem.2008.112797

Chang, K.H., Mestdagh, P., Vandesompele, J., Kerin, M.J., Miller, N., 2010. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. BMC Cancer 10. doi:10.1186/1471-2407-10-173

De Kok, J.B., Roelofs, R.W., Giesendorf, B.A., Pennings, J.L., Waas, E.T., Feuth, T., Swinkels, D.W., Span, P.N., 2005. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab.

Invest. 85, 154–159. doi:10.1038/labinvest.3700208

Floyd, S.K., Bowman, J.L., 2005. MicroRNAs: Micro-managing the Plant Genome, in: Meyer, P. (Ed.), Plant Epigenetics. Blackwell Publishing Ltd, Oxford, UK, pp. 244–278.

Frericks, M., Esser, C., 2008. A toolbox of novel murine house-keeping genes identified by meta-analysis of large scale gene expression profiles. Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1779, 830–837. doi:10.1016/j.bbagrm.2008.08.007

Glare, E.M., Divjak, M., Bailey, M.J., Walters, E.H., 2002. beta-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels. Thorax 57, 765–770.

Kozera, B., Rapacz, M., 2013. Reference genes in real-time PCR. J. Appl. Genet. 54, 391–406. doi:10.1007/s13353-013-0173-x

Kwon, M.J., Oh, E., Lee, S., Roh, M.R., Kim, S.E., Lee, Y., Choi, Y.-L., In, Y.-H., Park, T., Koh, S.S., Shin, Y.K., 2009. Identification of Novel Reference Genes Using Multiplatform Expression Data and Their Validation for Quantitative Gene Expression Analysis. PLoS ONE 4, e6162. doi:10.1371/journal.pone.0006162

Lee, S., Jo, M., Lee, J., Koh, S.S., Kim, S., 2007. Identification of novel universal housekeeping genes by statistical analysis of microarray data. J. Biochem. Mol. Biol. 40, 226–231.

Leinonen, R., Sugawara, H., Shumway, M., on behalf of the International Nucleotide Sequence Database Collaboration, 2011. The Sequence Read Archive. Nucleic Acids Res. 39, D19–D21. doi:10.1093/nar/gkq1019

Meng, F., Henson, R., Wehbe-Janek, H., Smith, H., Ueno, Y., Patel, T., 2007. The MicroRNA let-7a Modulates Interleukin-6-dependent STAT-3 Survival Signaling in Malignant Human Cholangiocytes. J. Biol. Chem. 282, 8256–8264. doi:10.1074/jbc.M607712200

Millar, A.A., Waterhouse, P.M., 2005. Plant and animal microRNAs: similarities and differences. Funct. Integr. Genomics 5, 129–135. doi: 10.1007/s10142-005-0145-2

Popovici, V., Goldstein, D.R., Antonov, J., Jaggi, R., Delorenzi, M., Wirapati, P., 2009. Selecting control genes for RT-QPCR using public microarray data. BMC Bioinformatics 10, 42. doi: 10.1186/1471-2105-10-42

Saviozzi, S., Cordero, F., Iacono, M.L., Novello, S., Giorgio, S.V., Calogero, R.A., 2006. Selection of suitable reference genes for accurate normalization of gene expression profile studies in non-small cell lung cancer. BMC Cancer 6. doi:10.1186/1471-2407-6-200

Stahlhut Espinosa, C.E., Slack, F.J., 2006. The role of microRNAs in cancer. Yale J. Biol. Med. 79, 131–140.

Zhan, C., Yan, L., Wang, L., Jiang, W., Zhang, Y., Xi, J., Chen, L., Jin, Y., Qiao, Y., Shi, Y., Wang, Q., 2014. Identification of reference miRNAs in human tumors by TCGA miRNA-seq data. Biochem. Biophys. Res. Commun. 453, 375–378. doi:10.1016/j.bbrc.2014.09.086.

How to cite this article:

Rajesh Kumar, Avinash Marwal and Gaur, R.K. 2017. Identification of Potential Reference miRNA for qRT-PCR Studies in Cancers Using miRNA-seq Data. *Int.J.Curr.Microbiol.App.Sci.* 6(12): 1327-1333. doi: https://doi.org/10.20546/ijcmas.2017.612.150